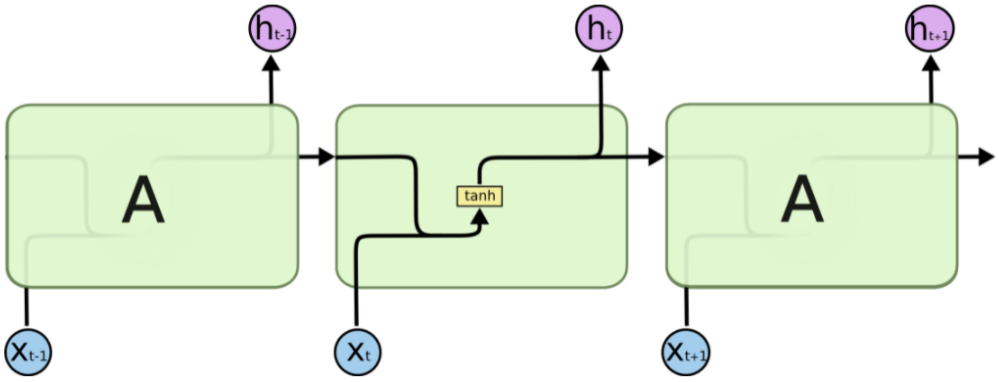
LSTM

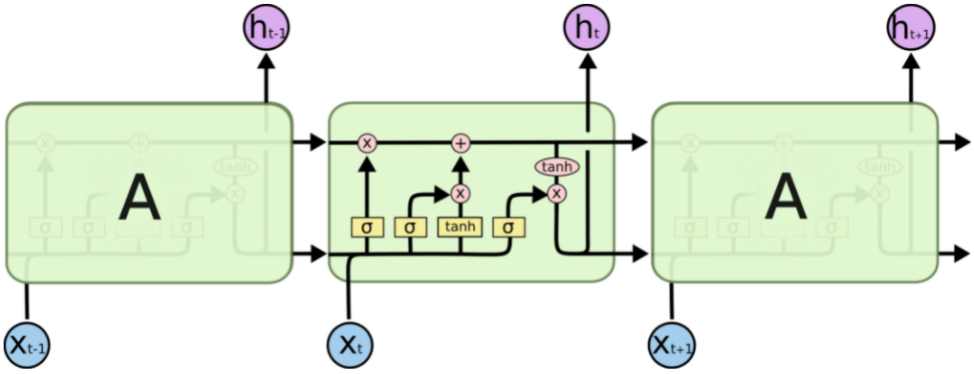
One of the advantages of RNNs over ANN is that they can link previous information to the current task. In such cases, where the gap between the relevant information and the place that it's needed is small, RNNs can learn to use the past information. But as that gap grows, RNNs become unable to learn to connect the information.

Long Short Term Memory networks (LSTMs) – are a special kind of RNN that are explicitly designed to avoid the long-term dependency problem.

In standard RNNs, this repeating module will have a very simple structure, such as a single tanh layer.



LSTMs also have this similar structure, but the repeating module has a different structure. Instead of having a single neural network layer, there are four, interacting in a very special way.



Useful Links to learn LSTM

- https://colah.github.io/posts/2015-08-Understanding-LSTMs/
- https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21
- https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/